National Institutes of Health
Office of Management

Bibliometrics for Portfolio Analysis

Chris Belter, Informationist, NIH Library
Christopher.Belter@nih.gov

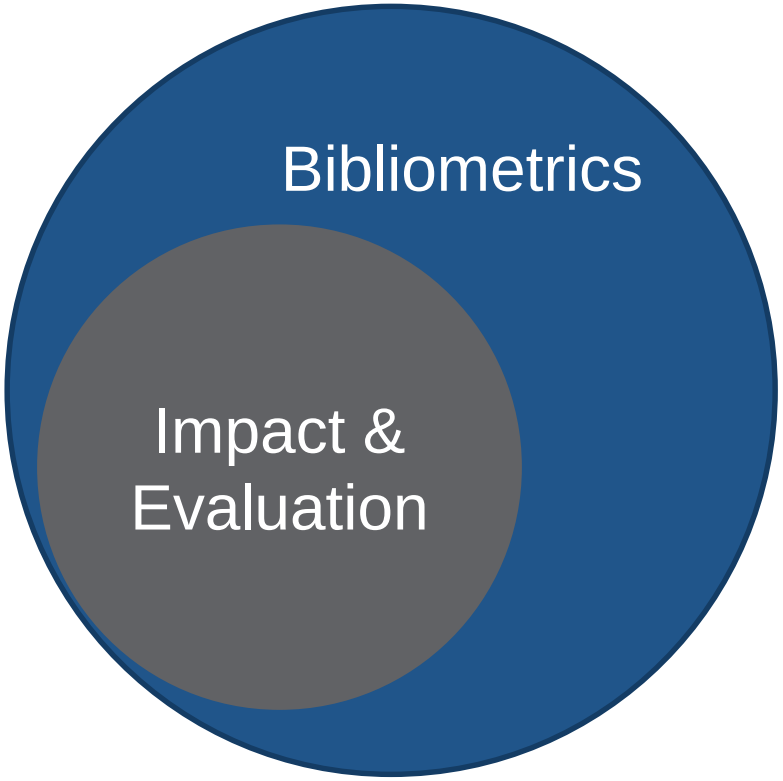
Office of Research Services
National Institutes of Health
U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

Agenda

- About bibliometrics
- Types of questions
  - Collaboration
  - Research topics
  - Citation Impact
- Data sources and tools

About Bibliometrics

What are Bibliometrics?



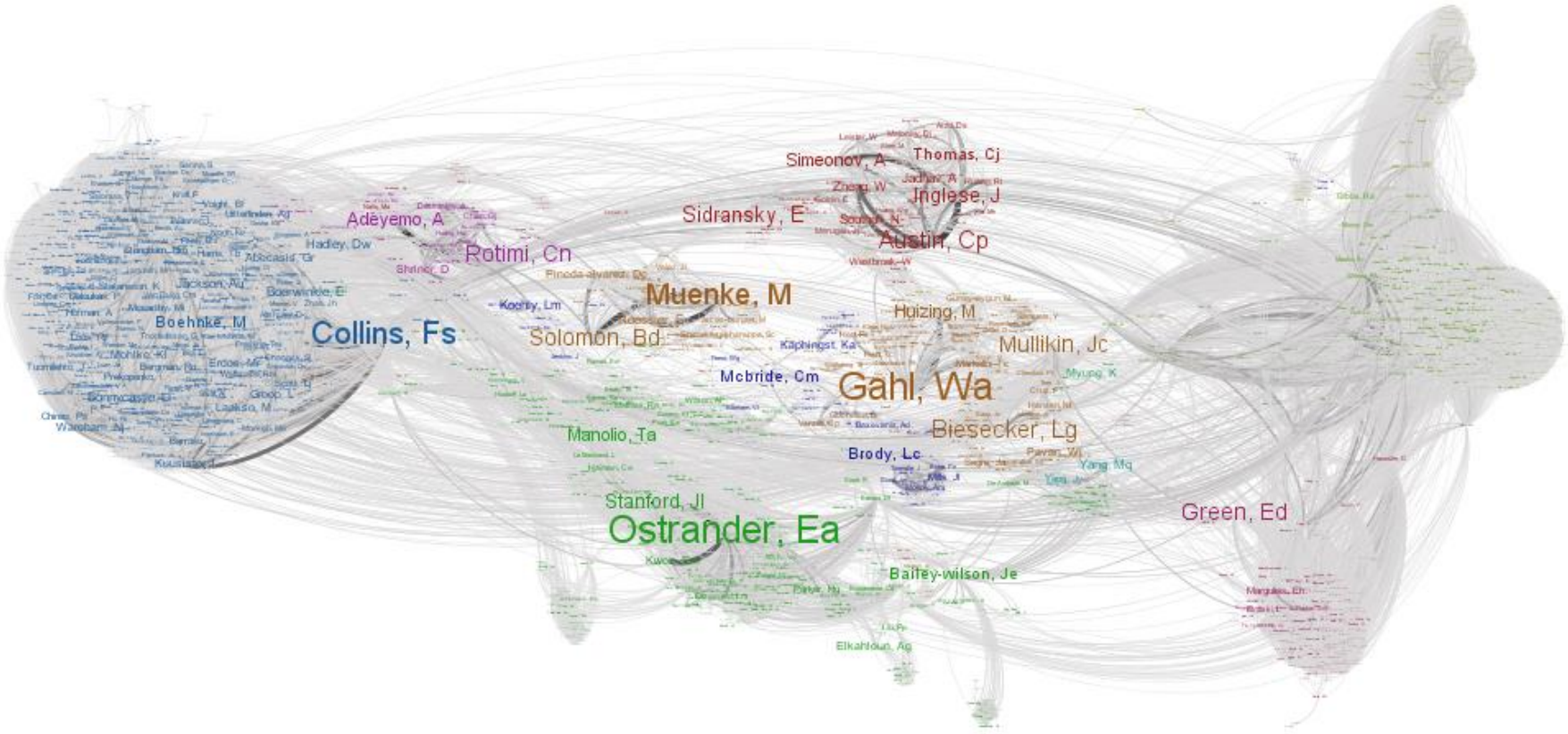
Bibliometrics are quantitative methods of studying scientific research using publications as a proxy for research

Bibliometrics: Pros and Cons

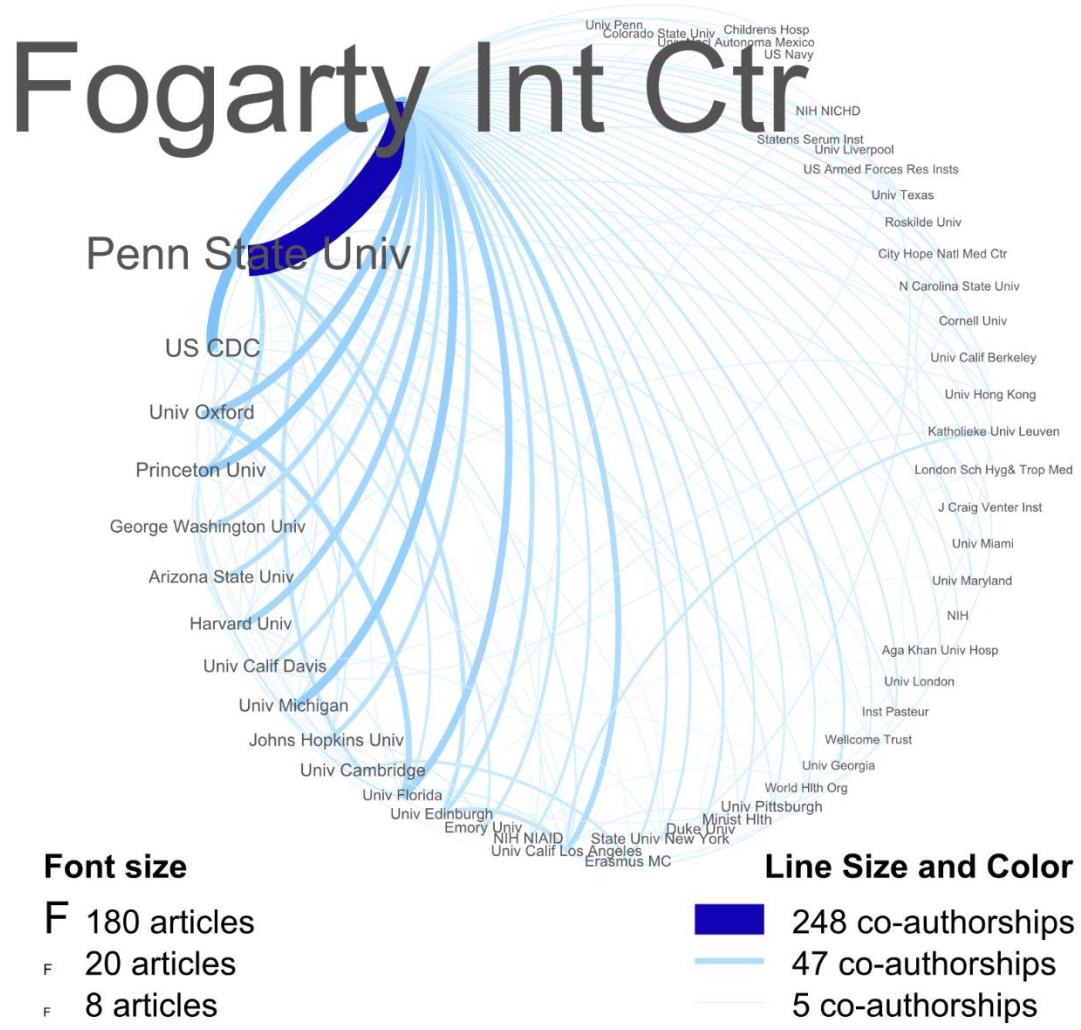
Pros	Cons
Handles large data sets	Requires expertise to generate and interpret
Produces reproducible results	Only measures publications
Impact according to a large sample	Limited measure of impact

Types of Questions: Collaboration

Individual Co-Authorship



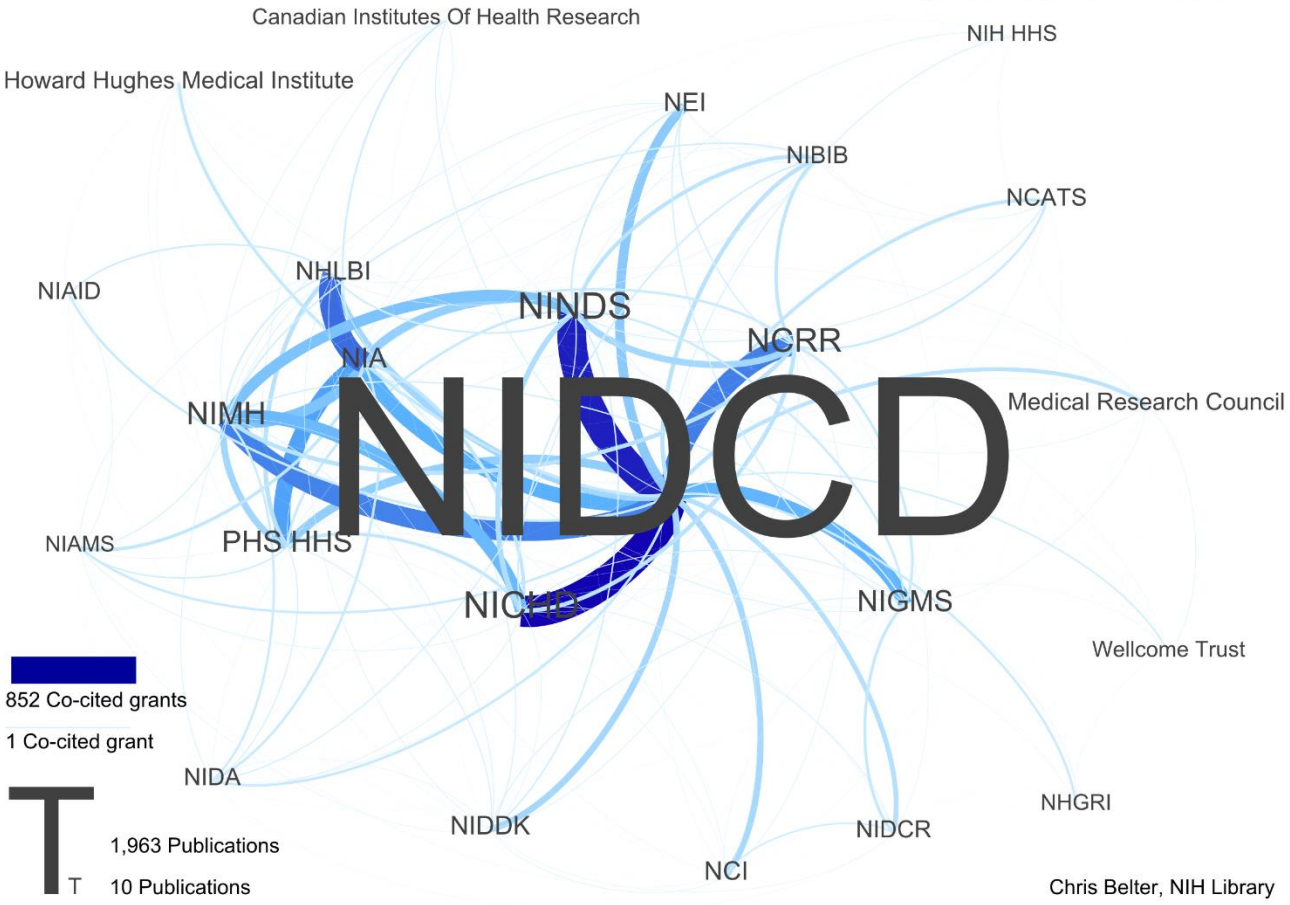
Institutional Co-Authorship



Institute Funding

NIDCD Co-Funding

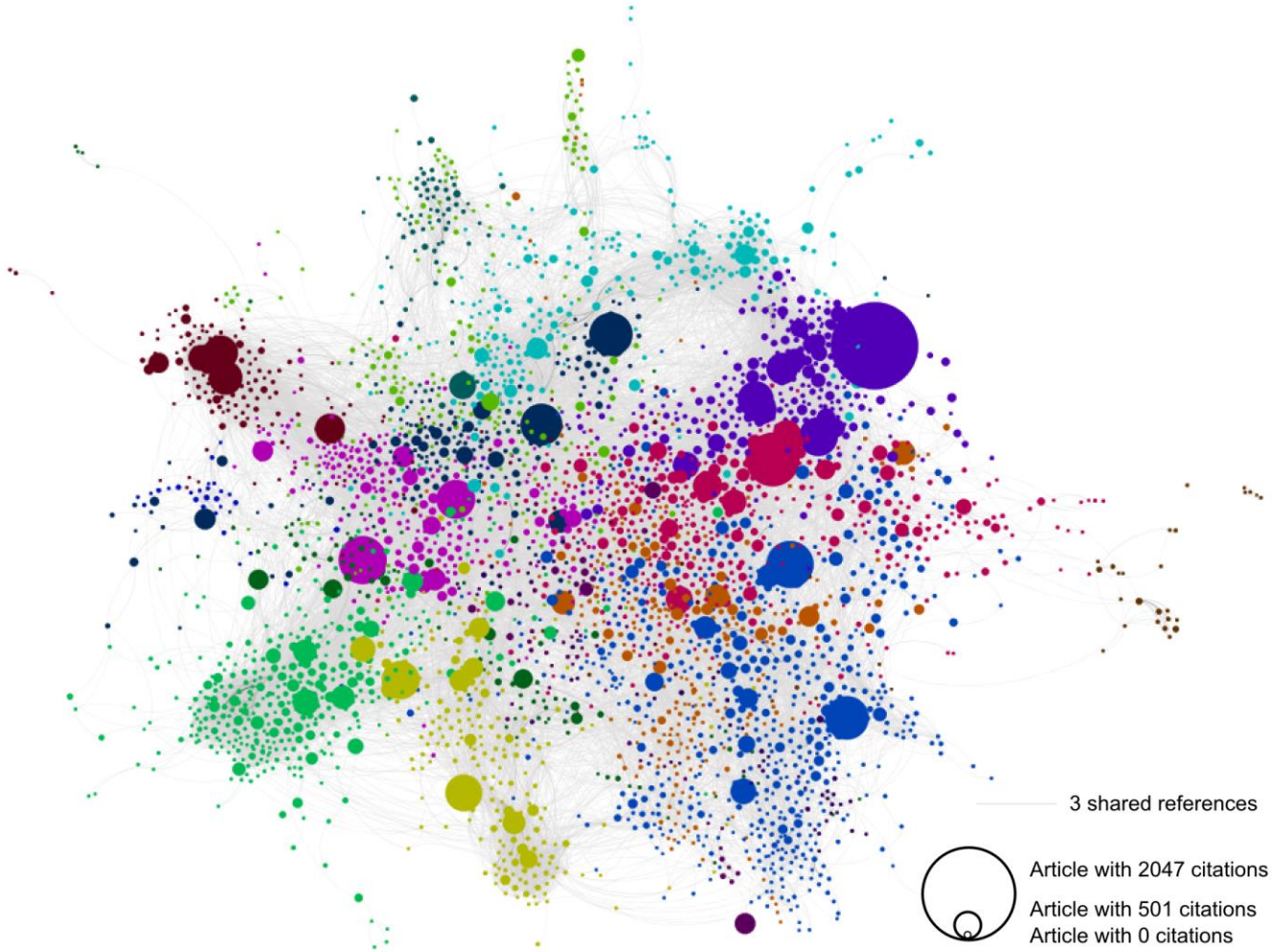
Based on grant numbers cited by publications in PubMed that were published during calendar year 2013. Only institutes that were co-cited with NIDCD by at least 10 publications are shown.



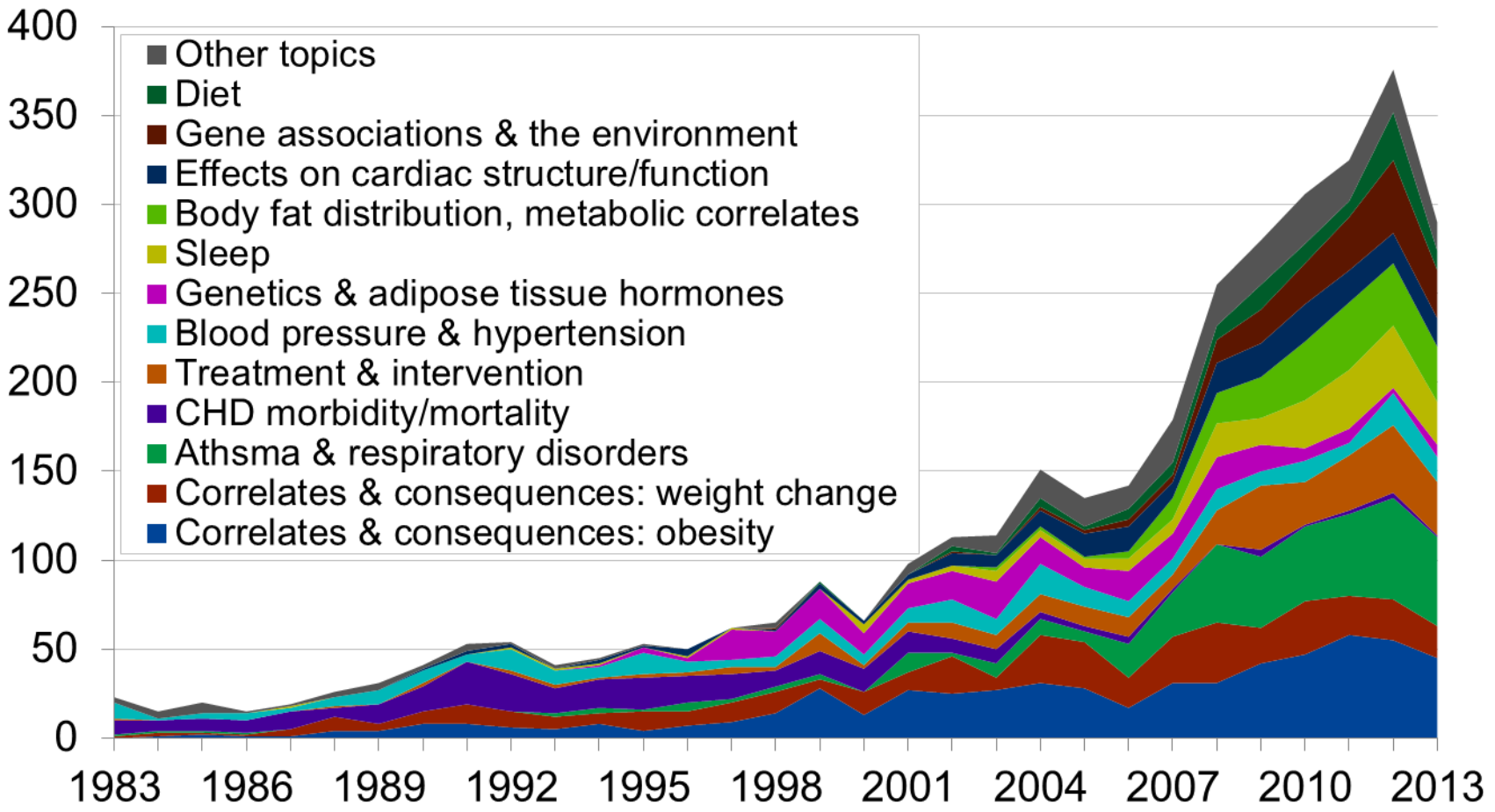
Chris Belter, NIH Library

Types of Questions: Research Topics

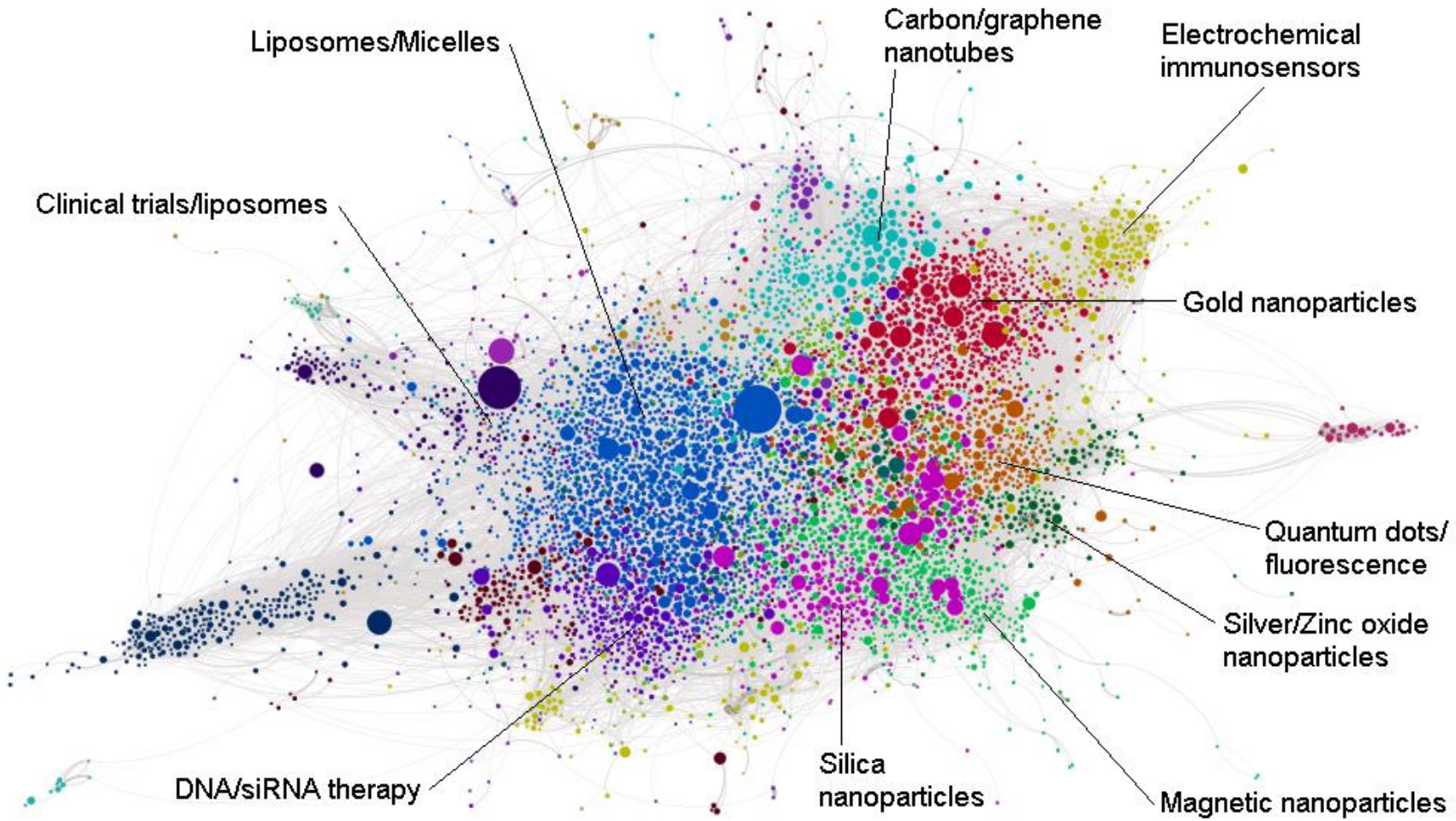
Paper Topics



Topic Evolution



Landscape Analysis



Types of Questions: Citation Impact

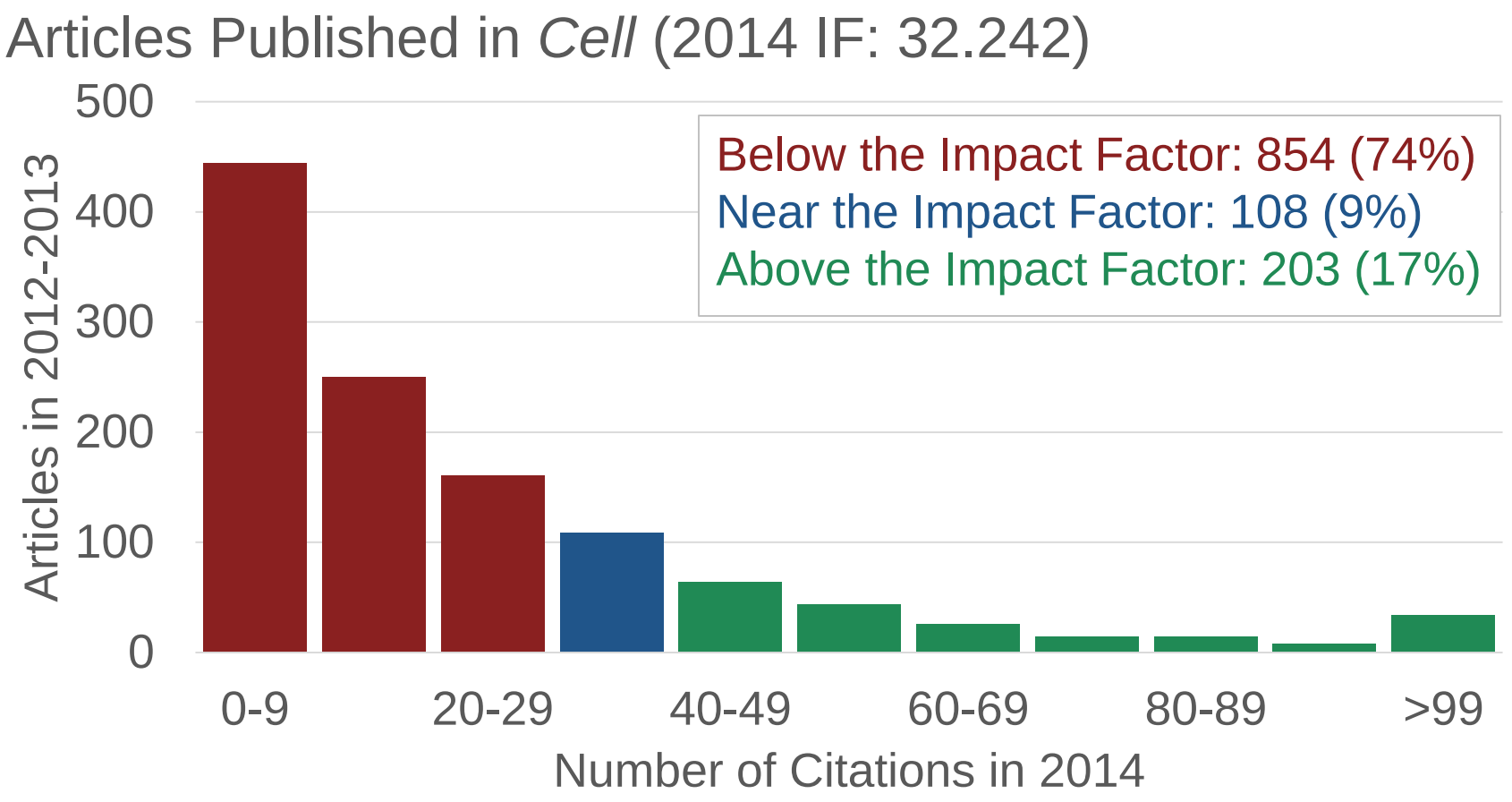
Citation-Based Indicators

Citations indicate the value, or usefulness, of one paper to the authors of another

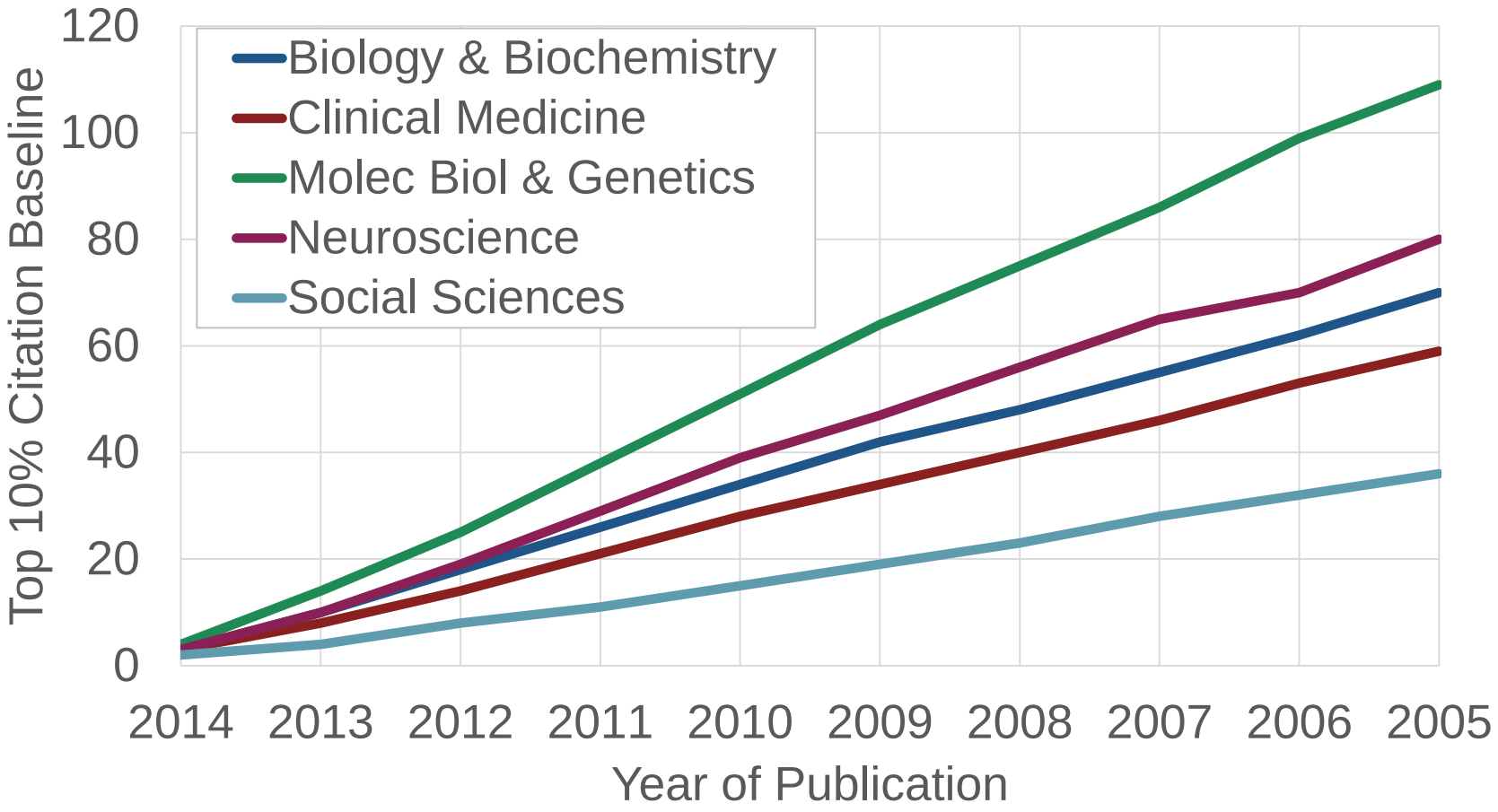
Measuring the total citations received by a paper indicates its usefulness to other authors

Measuring citations to a set of papers (by a laboratory, grant portfolio, institute, etc.) indicates the set's usefulness to science

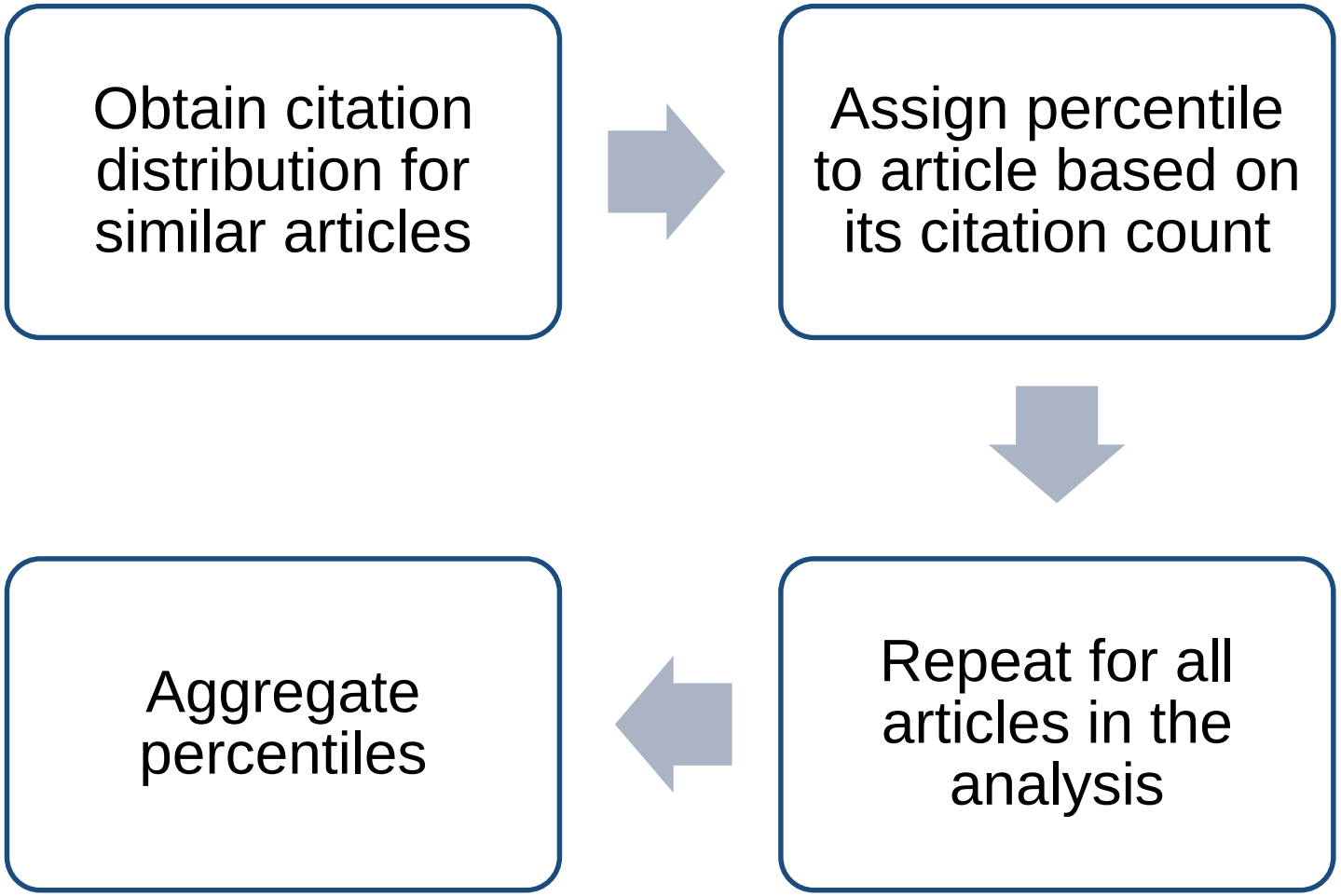
The Impact Factor



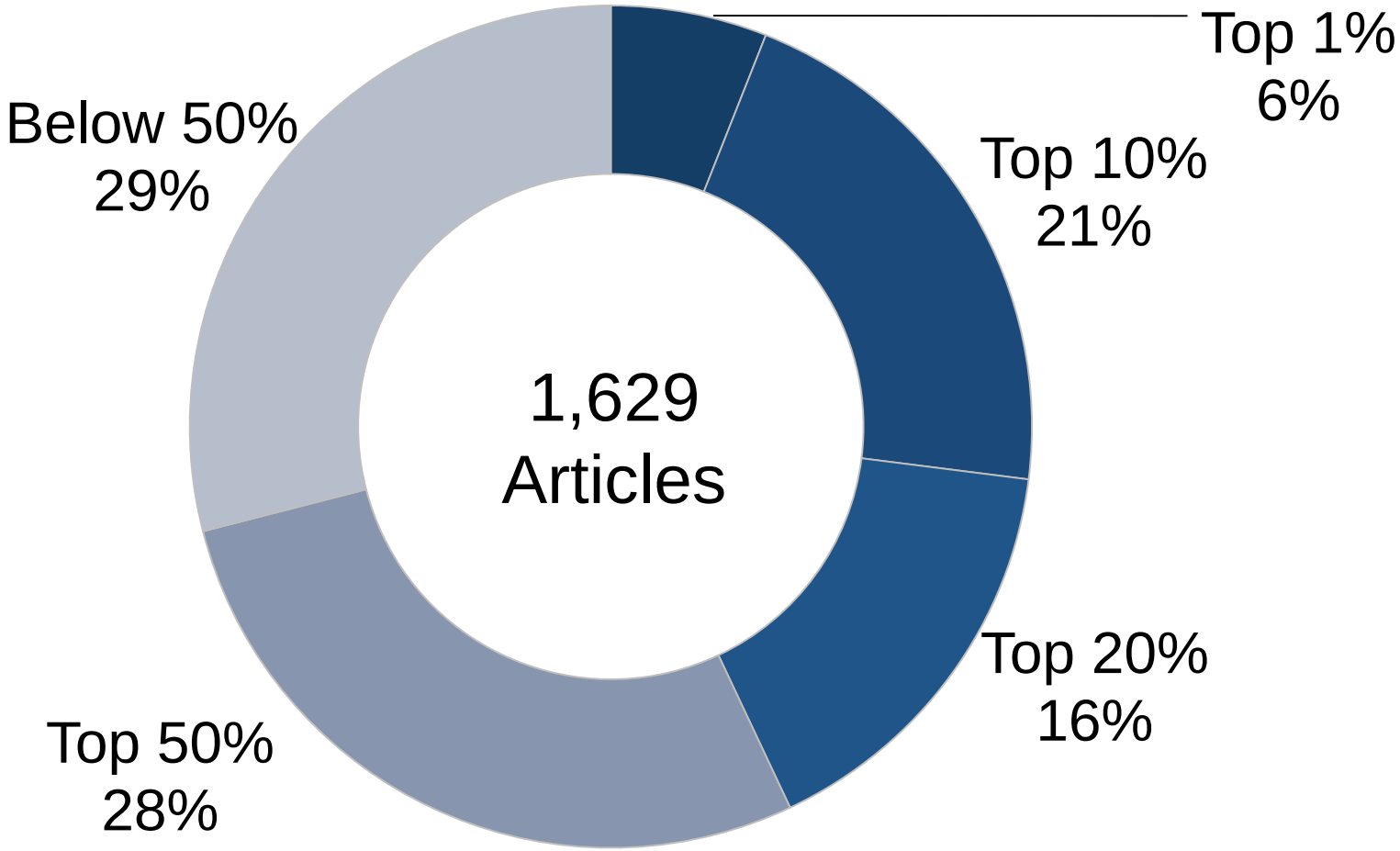
Age and Discipline



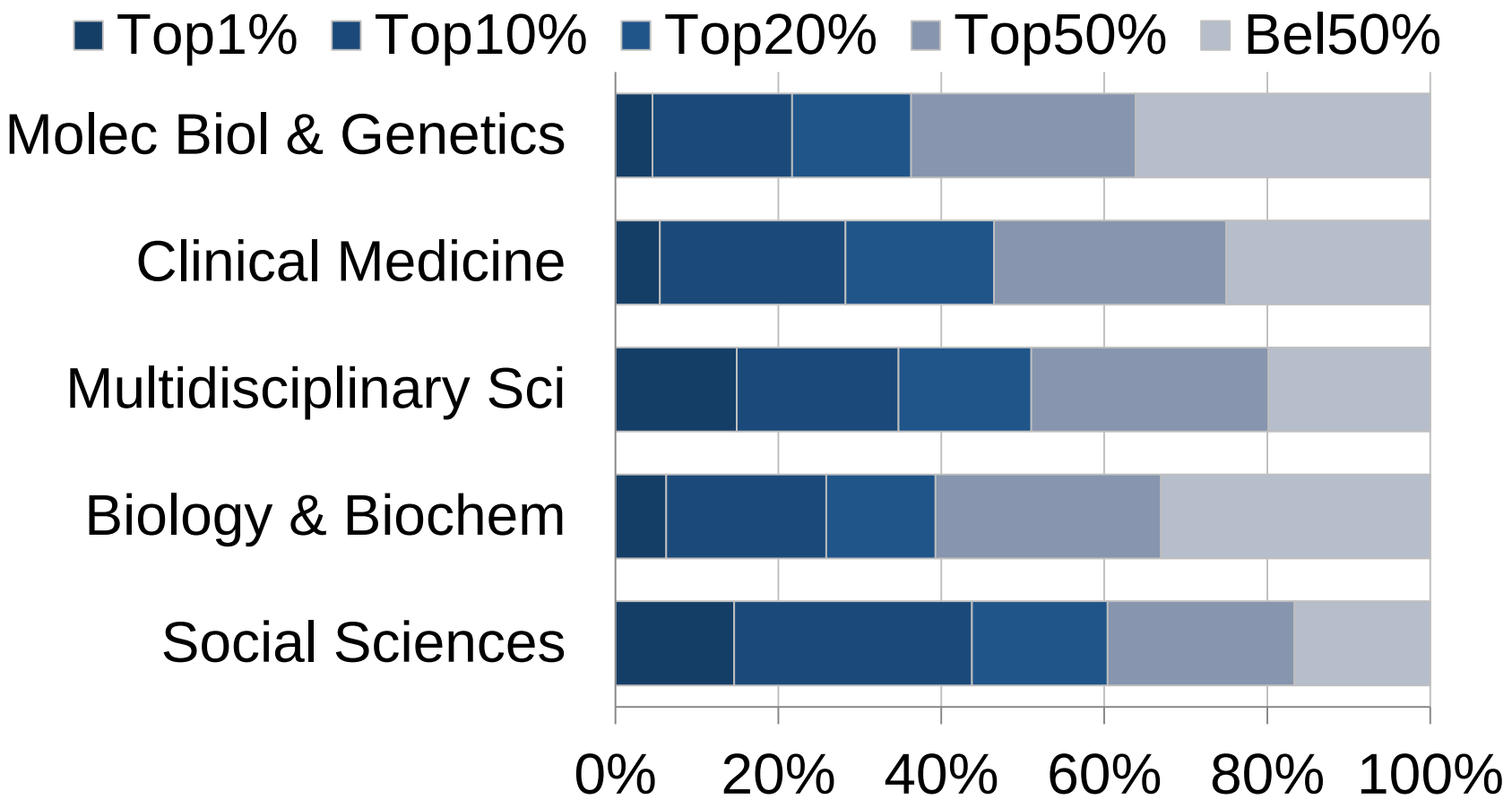
Percentile Ranking



Rank Distribution



Rankings by Discipline



Data Sources and Tools

Data Sources

Publication Data	• QVR, RePORTER, PubMed, Web of Science, Scopus, etc.
Analysis Data	• PubMed, Web of Science, Essential Science Indicators, Scopus, etc.

Analysis Tools

Data cleaning	• Google Refine and Gephi
Network analysis	• Science of Science (Sci2) Tool
Network visualization	• Gephi
Text mining	• Sci2 and VantagePoint (and R)

Questions?

Chris Belter
Christopher.Belter@nih.gov
301.451.5861